Plant Physiol. Biochem. 39 (2001) 295–311

S0981942800012377/FLA

A genomics approach towards salt stress tolerance

Hans J. Bohnert[a,b]*, Patricia Ayoubi[d§], Chris Borchert[a§], Ray A. Bressan[c], Robert L. Burnap[d], John C. Cushman[e], Mary Ann Cushman[e§], Michael Deyholos[b§], Robert Fischer[a], David W. Galbraith[b], Paul M. Hasegawa[c], Matt Jenks[c§], Shinji Kawasaki[a§], Hisashi Koiwa[c§], Shin Kore-eda[e§], Byeong-Ha Lee[b§], Chris B. Michalowski[a§], Eduardo Misawa[d], Mika Nomura[e§], Neslihan Ozturk[a§], Bradley Postier[d§], Rolf Prade[d], Chun-Peng Song[b§], Yuko Tanaka[b§], Hong Wang[b§], Jian-Kang Zhu[b]

[a] Department of Biochemistry, University of Arizona, Tucson, AZ 85721, USA

[b] Department of Plant Sciences, University of Arizona, Tucson, AZ 85721, USA

[c] Horticulture Department, Purdue University, West Lafayette, IN 47907, USA

[d] Department of Microbiology & Molecular Genetics, Oklahoma State University, Stillwater, OK 74078, USA

[e] Department of Biochemistry, University of Nevada, Reno, NV 89557, USA

Received 11 October 2000; accepted 11 November 2000

Abstract – Abiotic stresses reduce plant productivity. We focus on gene expression analysis following exposure of plants to high salinity, using salt-shock experiments to mimic stresses that affect hydration and ion homeostasis. The approach includes parallel molecular and genetic experimentation. Comparative analysis is employed to identify functional isoforms and genetic orthologs of stress-regulated genes common to cyanobacteria, fungi, algae and higher plants. We analyze global gene expression profiles monitored under salt stress conditions through abundance profiles in several species: in the cyanobacterium *Synechocystis* PCC6803, in unicellular (*Saccharomyces cerevisiae*) and multicellular (*Aspergillus nidulans*) fungi, the eukaryotic alga *Dunaliella salina*, the halophytic land plant *Mesembryanthemum crystallinum*, the glycophytic *Oryza sativa* and the genetic model *Arabidopsis thaliana*. Expanding the gene count, stress brings about a significant increase of transcripts for which no function is known. Also, we generate insertional mutants that affect stress tolerance in several organisms. More than 400 000 T-DNA tagged lines of *A. thaliana* have been generated, and lines with altered salt stress responses have been obtained. Integration of these approaches defines stress phenotypes, catalogs of transcripts and a global representation of gene expression induced by salt stress. Determining evolutionary relationships among these genes, mutants and transcription profiles will provide categories and gene clusters, which reveal ubiquitous cellular aspects of salinity tolerance and unique solutions in multicellular species. © 2001 Éditions scientifiques et médicales Elsevier SAS

Arabidopsis T-DNA tag / EST sequencing / functional genomics / gene expression profiles / microarray analysis / phylogenetic profiles of stress tolerance / salinity stress / stress tolerance

EST, expressed sequence tag / ORF, open reading frame

1. INTRODUCTION

For at least 8 000 years, mankind has manipulated plant species from subtropical regions to grow under diverse climatic conditions. The results of this activity provide all our food, feed and fiber. Domestication extended to just a few hundred species, with maybe ten of these providing the majority of all foods. Selection, increasingly based on scientific criteria, has allowed us to keep pace with population growth [9]. Population growth, changes in lifestyle, competition for fresh water between farmers and cities, and possible global environmental changes have led to alarmist projections that seem to argue for additional strategies by which food supply can be guaranteed [32]. Areas where food production is problematic coincide largely

*Correspondence and reprints: fax +1 520 621 1697.
E-mail address: bohnerth@email.arizona.edu (H.J. Bohnert).

§ Contributed experimental data from ongoing sub-projects within the stress genomics consortium.

with places where water is a precious commodity, prolonged droughts are frequent, where saline and marginal soils are found and where temperature extremes compromise production. Moreover, the increased productivity achieved in irrigated agriculture has become a double-edged sword, because salinization following prolonged irrigation is unavoidable [13, 37]. These considerations have galvanized strong interest in studying plant abiotic stress responses and understanding the meaning of stress tolerance as a biological phenomenon.

Understanding abiotic stress tolerance, not to mention breeding for stress tolerance, proved difficult because of the trait's multigenicity. As a consequence no traditional crop lines exist that combine tolerance to high salinity or drought with high yield, which represents yet another complex trait. Many physiological studies have analyzed the problem. In the search for tolerance mechanisms, plant biologists have generated hypotheses by interpreting correlative evidence from many species based on biochemical and biophysical principles that govern stress tolerance or resistance [27] These insights provided guiding principles for moving from physiology to protein and enzyme analysis, genetic structure, gene function and gene expression studies, finally culminating in transgenic and mutant generation and analyses.

The last decade has brought a sea change in our views about stress sensitivity fueled by results from molecular genetics. Changes in the expression of individual genes and proteins induced by stress have been monitored under different conditions. As of the year 2000, the sequence of the *Arabidopsis thaliana* genome is nearly completed, and soon we will have a catalog of plant gene expression exceeding a million transcripts (http://www.ncbi.nlm.nih.gov/dbEST/index.html) [28, 44, 45]. While numbers are important, more important is that from the molecular analysis of these transcripts, collected as ESTs (expressed sequence tags), the frequency and type of transcribed genes are obtained for organs, tissues and cells during development and under various perturbations. Such expression profiles are augmented by the analysis of EST-derived microarrays or oligonucleotide-based DNA-chips both providing information on changes in the expression levels over time or during experimental perturbations of the plants [7, 39, 50]. Accompanying these activities has been yet another breakthrough, whose impact on stress biology and crop breeding will be, we think, even more profound. In *A. thaliana*, rice and corn saturation mutagenesis seems to have been achieved [4, 31, 36]. The resulting mutants make possible proof of concept experiments on a large scale.

Here we present an initial set of data from an ongoing project that is focussing on defining the complete gene set associated with salinity stress. In addition, we report on Arabidopsis mutants with altered stress sensing and signal transduction characteristics. New technologies are employed to advance understanding: large-scale EST preparation and expression profiling, microarray analysis identifying target genes whose expression is stress-regulated, and the generation of mutants in stress-relevant signal transduction and response pathways. As the first results emerge, additional strategies can be employed that target the analysis of the many functionally unknown genes or transcripts, which are regulated when plants experience abiotic stress.

2. RESULTS AND DISCUSSION

For continued vegetative growth and the development of reproductive organs under osmotic and ionic stress conditions, plants must, above all, obtain water. To satisfy this essential need, each of many adaptive mechanisms that have been detected as plant defenses must be subordinate to this goal. When stomata are closed to limit water loss, a series of events adjusts photosynthesis, carbon fixation and carbohydrate transport, initiating processes that maintain the integrity of the photosynthetic and carbon fixation apparatus [18, 29]. The complex protective mechanisms include redox control, scavenging of radical oxygen species, metabolite accumulation and altered partitioning, and ion homeostasis [17]. Testing of physiological principles has generated many data and some correlations which have in part been tested by transgenic alterations and further physiological studies [5, 16, 17, 21]. As minuscule as the stress-ameliorating effects have been resulting from this first generation of transgenic models, they have highlighted mechanisms. Large-scale EST and expression analysis can now be used to scrutinize, reinforce or replace the results from earlier studies. One outcome seems already clear. While the presence of the appropriate synthetic machinery is essential for tolerance acquisition, it is very likely that many or most of the necessary biochemical elements are present in all species. What distinguishes species, it seems, is: how fast and how persistent this machinery is engaged; how stress is perceived; and how signals are subsequently integrated into a network of responses [17, 33]. Providing examples, we discuss strategies designed

to answer three questions, which we consider important to understand what osmotic/ionic stress tolerance entails. First, a view across phyla and orders on the nature of genes that are regulated during stress provides a spectrum of mechanisms for cell-specific determinants of tolerance. We approach this by generating cDNA libraries for the stressed state in several model organisms, transcript abundance profiles and microarray analyses of stress-regulated transcripts. A second objective is determining how signaling networks control adaptive mechanisms. The generation and analysis of tagged mutants in Arabidopsis let us approach this aspect. The third way is long-term. In order to understand stress tolerance, functions must be determined for the many stress-regulated transcripts and genes that have not yet been characterized and studied. For this purpose, information is needed about where, when and to what extent functionally unknown transcripts and proteins are expressed and what these functions are.

2.1. Transcripts from salt-stressed plants and changes in transcript abundance during stress

Among the cDNA libraries that have been established for many plant species, very few have been generated with tissues from stressed plants. We use nearly exclusively cDNA libraries generated with RNA from stressed rice (*Oryza sativa* L., Pokkali), ice plant (*Mesembryanthemum crystallinum* L.) and *Arabidopsis* (Columbia). In addition, libraries and ESTs were generated for salt-stressed corn, barley, tobacco, and from the alga *Dunaliella salina* during an upshift in the medium salt concentration. Libraries are available from various tissues, mainly roots and leaves, at different developmental stages and at different times during stress (*table I*). Classification of the ESTs into functional categories provided changes, relative to ESTs from unstressed organisms, associated with high salinity. In salt-stressed rice, for example, the categories ‘functionally unknown’, ‘transport facilitation’, ‘cell rescue and defense’ increase strongly. In the categories ‘energy’ and ‘metabolism’, with small overall increases, the nature of the transcripts, compared to the unstressed state, changes. Among the functions of energy provision, transcripts relating to mitochondrial functions increase as do transcripts for functions in secondary metabolism (Kawasaki S., Deyholos M., Borchert C., Brazille S., Kawai K., Galbraith D.W., Bohnert H.J., unpubl.). *Figure 1* and *tables II* and *III* provide a transcript abundance profile for roots from salt-stressed corn (*Z. mays* L.) and rice leaves annotated in general functional categories.

Progress and success of large-scale DNA sequencing and analysis projects depend crucially on the availability of suitable bioinformatics programs. EST sequences, obtained from mRNA populations from various plant and genomic model systems and tags selected through random sampling, were analyzed using a program designed to accommodate the stress tolerance focus of the consortium. A total of 30 783 EST sequences were processed and analyzed using PipeOnline_v2.0b (Ayoubi P., Jin X., Leite S., Liu X., Martajaja J., Abduraham R., Wan Q., Yan W., Misawa E., Prade R.A., unpubl.), resulting in 16 090 unique sequences (*table IV*). PipeOnline is a flexible, automated web-interactive tool offering processing of large batches of raw DNA sequence data with functional classification of putative genes. Like all annotation tools, PipeOnline is prone to misalignments and must be considered as a tool whose characteristics still evolve. PipeOnline produces a portable database containing the edited, assembled nucleotide sequences, the BLASTX results of the closest homologs and provides an assessment of probable function based on automated functional assignment algorithms. This functional assignment algorithm within PipeOnline utilizing the information generated by BLASTX homology analysis is for estimation of function using MPW-based functional overview [40, 41] providing a preliminary metabolic overview for raw DNA sequence data derived from large-scale sequencing projects.

Obvious is a significant increase of ESTs for functionally unknown sequences with homology to ESTs or genes from other organisms, many from non-plant species. Most striking, however, is the increase of sequences that produced no hits, i.e. no meaningful homology by BLASTX criteria to sequences in either the non-redundant or EST databases. In an analysis of approximately 2 500 ESTs from salt-stressed rice (150 mM, line Pokkali) for example, the number of such ‘no hit’ annotations increased from 8.5 % in the unstressed state to 16 % in cDNAs from stressed plants (Kawasaki S., Deyholos M., Borchert C., Brazille S., Kawai K., Galbraith D.W., Bohnert H.J, unpubl.). Similar increases characterize the stressed state in *Arabidopsis* and *Mesembryanthemum* (Cushman J.C., Nomura M., Kore-eda S., Bohnert H.J., Deyholos M., Michalowski C.B., Hasegawa P.M., unpubl. data). For example, the analysis of nearly 4 000 randomly sampled ESTs from *M. crystallinum* showed a ‘no hit’ annotation in ESTs from unstressed plants at 13 %, which increased to 17 % in salt-

Table I. cDNA libraries for salinity stress-relevant ESTs. * Expressed as mM NaCl and length of stress treatment. For *D. salina*, molar concentrations of NaCl were used.

Name	Organism	Tissue	Stress*	Plant age
AA	*A. thaliana* (Columbia)	Leaves	200; 20 h	12 weeks
AB	*A. thaliana* (Columbia)	Leaves	200; 3, 6, 9, 12 h	2–3 weeks
AC	*A. thaliana* (Columbia)	Leaves and roots	150; 20 h	2–3 weeks
AD	*A. thaliana* (Columbia)	Seedlings	Subtracted, 160 mM NaCl, 4 h	10–14 d
AF	*A. thaliana* (Columbia)	Roots	200; 6 h	Flowering
DA	*D. salina*	Cells	shift 1.7 and 2.5 M to 3.4 M; 5 h	Cell culture
HA	*H. vulgare* (Tokak)	Leaves	No stress	~3 weeks
HB	*H. vulgare* (Tokak)	Leaves	Drought; 6 and 10 h	~3 weeks
HC	*H. vulgare* (Tokak)	Roots	Drought; 6 and 10 h	~3 weeks
HD	*H. vulgare* (Tokak)	Roots	No stress	~3 weeks
MA	*M. crystallinum*	Leaves	No stress	5–6 weeks
MB	*M. crystallinum*	Leaves	400; 30 h	5–6 weeks
MR	*M. crystallinum*	Leaves	400; 48 h	5–6 weeks
MC	*M. crystallinum*	Roots	No stress	5–6 weeks
ME	*M. crystallinum*	Roots	400; 6 h	5–6 weeks
MF	*M. crystallinum*	Roots	400; 12 h	5–6 weeks
MG	*M. crystallinum*	Roots	400; 30 h	5–6 weeks
MH	*M. crystallinum*	Roots	400; 78 h	5–6 weeks
MI	*M. crystallinum*	Seedlings	250; 3 d	14 d
MJ	*M. crystallinum*	2° Leaves	500; 5 d	8–9 weeks
MK	*M. crystallinum*	1° Leaves	500; 6–7 d	6 weeks
ML	*M. crystallinum*	Flowers and seedpods	500; 6 weeks	> 12 weeks
MM	*M. crystallinum*	Bladder cells	500; 6 weeks	> 12 weeks
MN	*M. crystallinum*	Side shoots	500; 3 d	6 weeks
MO	*M. crystallinum*	Meristems	No stress	5 weeks
MP	*M. crystallinum*	Meristems	500; 3 d	6 weeks
NA	*N. tabacum* (SR1)	Roots and leaves	200; 6, 24, 30, 48 h	4 weeks
NB	*N. tabacum* (SR1)	Leaves	No stress	~10 weeks
OA	*O. sativa* (Nipponbare)	Root	200; 19 h	3–4 weeks
OB	*O. sativa* (Nipponbare)	Leaves	200; 19 h	3–4 weeks
OC	*O. sativa* (Pokkali)	Roots	No stress	1 week
OD	*O. sativa* (Pokkali)	Roots	150; 1 d	1 week
OE	*O. sativa* (Pokkali)	Roots	150; 2, 3 d	1 week
OF	*O. sativa* (Pokkali)	Roots	150; 1 week	2 weeks
OG	*O. sativa* (Pokkali)	Leaves	150; 1, 2, 3 d, 1 week	1 week
OH	*O. sativa* (Pokkali)	Leaves	No stress	1–2 weeks
ZA	*Z. mays* (B73)	Roots	150; 24 h	2 weeks
ZB	*Z. mays* (B73)	Leaves and shoots	150; 24 h	2 weeks

stressed plants (500 mM). Also, *M. crystallinum* ESTs with significant matches to genes with unknown or unclear functions increased from 33 % in well-watered plants to 47 % following salt stress (Cushman M.A., Dennis M., Bufford D., Akselrod I., Landrith D., Maroco J., Kore-eda, S., Rogers S. Cushman J.C., unpubl.).

Global transcript abundance profiles are only one criterion to distinguish stressed and unstressed plants in functional categories of genes whose expression can be expected to change. Within categories, differences can be observed in the expression of isoforms. Examples are the replacement of isoforms for glutathione-S-transferases, SAM-synthetases, various chaperones, or ascorbate peroxidases and other radical oxygen scavenging enzymes, possibly indicating that the same function is carried out by differentially regulated homologs. Thus, distinct isoforms for otherwise identical functions may be expressed based on promoter strength or regulation and tissue specificity. Protein stability or modification could also be different in the stress-induced isoforms but biochemical analyses on distinctive characters of such isoforms are missing. However, such replacement of one isoform by another is not always obvious, due to the limited DNA sequencing that is typical for high-throughput EST projects and also due to the inconsistencies of previous annotations, which can lead to the perpetuation of errors.

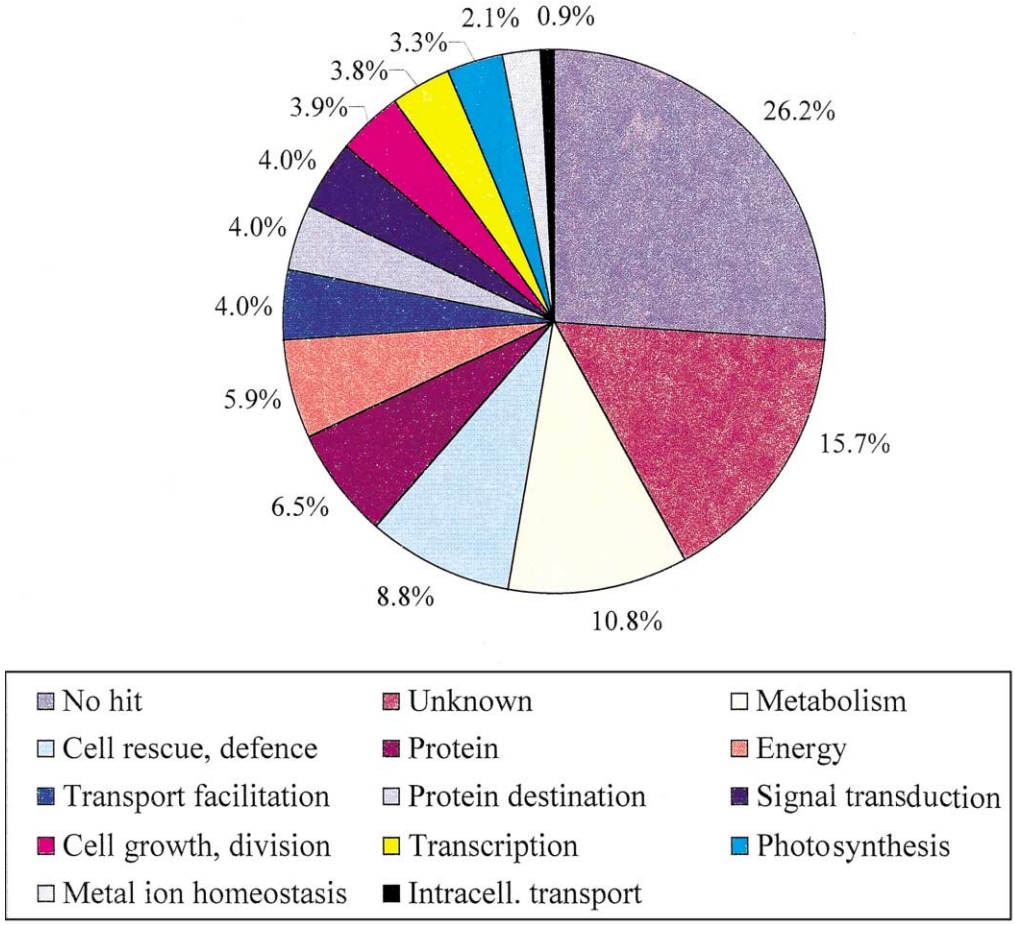

Figure 1. Chart of ESTs categorized by function. A total of 1 882 ESTs from salt stress roots (cDNA library ZA, *table I*) are included. Sequences were determined by the maize EST consortium and deposited at http://www.zmdb.iastate.edu.

This argument is also valid for microarray analyses because most elements deposited on the slides are derived from ESTs. Only sequencing of the 3'-ends of cDNAs and the deposition of precisely tailored 3'-ends and coding regions can reliably distinguish orthologs. Despite these caveats, transcript abundance profiles can lead to unexpected insights which validate the use of the technology. For example, most photosynthesis-related ESTs show a precipitous decline in expression during the initial hours of salt stress, yet the abundance of a few transcripts increase (Michalowski C.B., unpubl.). They represent specific isoforms for protein components of the reaction centers, light-harvesting and water-splitting complexes. In general, however, a strong decline in transcripts in the categories of photosynthesis and in cell growth and division is a typical stress response (Cushman M.A., Dennis M., Bufford D., Akselrod I., Landrith D., Maroco J., Kore-eda, S., Rogers S. Cushman J.C., unpubl.; Kawasaki S., Deyholos M., Borchert C., Brazille S., Kawai K., Galbraith D.W., Bohnert H.J., unpubl.).

2.2. Microarray analyses

Unless all genes from an organism are available, microarray analyses will remain incomplete, but even partial sets of arrayed ESTs, transcripts or genes can provide useful information. First, microarray analysis can be instrumental for studying tissue and cell specificity of transcript expression. *Table V* compares ice plant sequences in a microarray hybridization that are highly represented either in shoots or in epidermal bladder cells (EBC), which represent a sodium storage tissue in salinity-stressed ice plants [1]. It is not surprising that the EBC profile is characterized by

Table II. Abundance profile of ESTs in a root cDNA library of salt-stressed corn.

Gene category	No. of ESTs	% of total	Category description
No hit	493	26.2	BLAST search in non-redundant GenBank finds no homologs
Unknown	296	15.7	BLAST search GenBank identifies unknown, hypothetical, and putative proteins, or proteins with unknown functions
Metabolism	203	10.8	Amino acid, nucleotide, C-compound and carbohydrate, lipid, fatty acid and isoprenoid, nitrogen and sulfur, and secondary metabolisms
Cell rescue, defense	165	8.8	Environmental stimuli responses (stress, wounding, phytohormone regulation, and ion deficiency), radical scavenging, detoxification, DNA repair, and cell death
Protein synthesis	123	6.5	Ribosomal proteins, translation (initiation, elongation and termination), translational control, and tRNA synthetases
Energy	111	5.9	Glycolysis and gluconeogenesis, pentose-phosphate pathway, TCA pathway, respiration, metabolism of energy reserves (e.g. glycogen, trehalose), and fatty acid oxidation
Transport facilitation	76	4.0	Ion channels, water channels, ion, sugar and carbohydrate, amino acid, lipid, purine and pyrimidine transporters, transport ATPases
Protein destination	76	4.0	Protein folding and stabilization, protein targeting, sorting and translocation, protein modification (e.g. myristylation, farnesylation, palmityolation, glycosylation) and proteolysis
Signal transduction	75	4.0	Receptor proteins, second messenger such as calmodulins, key kinases, key phosphatases, and G-proteins
Cell growth, division	74	3.9	Cell growth, development, cell cycle control and mitosis, cytokinesis, and DNA synthesis and replication
Transcription	71	3.8	mRNA synthesis (including general transcription activities, transcriptional control such as transcriptional factors, and chromatin modification), mRNA splicing, and mRNA stabilization and degradation
Photosynthesis	63	3.3	Chlorophyll *a*/*b*-binding protein, ferredoxin related genes, photosystem I and II reaction center
Metal ion homeostasis	39	2.1	Homeostasis of metal ions and other ions
Intracellular transport	17	0.9	Vesicular transport (Golgi network), vacuolar transport, cytoskeleton-dependent transport
Total	1 882	100	

sequences that can be expected in an epidermis tissue: for example, ESTs coding for pathogen response proteins, antifungal proteins and lipid transfer proteins.

Figure 2 exemplifies one type of data resulting from microarray experiments. An array including approximately 2 600 elements from salt-stressed ice plants was hybridized with Cy3-labeled RNA from stems and Cy5-labeled RNA from epidermal bladder cells (EBC). The graph simultaneously compares hybridization intensity, indicating mRNA abundance, and tissue/cell specificity. In some highlighted elements (see *figure 2*) closely related (~83 %) protein kinases can thus be assigned to different cell types. The graph also shows transcripts for water channel proteins nearly exclusively expressed in the stems. In addition, the intensity distribution places many EBC-specific transcripts in the high-abundance category of transcripts, while stem-specific transcripts are of medium abundance. Transcripts for which no hits can be detected in the non-redundant and EST databases are distributed throughout the graph.

The following provides another example for the type of information gained by microarray analysis of ESTs from cDNA libraries derived from stressed tissues. Stress-inducible ESTs are more highly represented in libraries generated from tissues of stressed plants than in libraries from well-maintained plants, and stress-repressed ESTs are less highly represented in libraries from stressed plants. This was amply documented by the analysis of an Arabidopsis microarray in which 6 236 ESTs from the collection of Michigan State University (MSU libraries) were spotted together with 2 976 ESTs from salt-stressed plants (UA libraries) (Deyholos M., unpubl.). In this example, 67 ESTs from the MSU libraries were upregulated (67/6 236 = 1.1 %), while 77 ESTs from the UA libraries of stressed Arabidopsis were upregulated (77/2 976 = 2.6 %). A similar result characterized repressed cDNAs: 37 ESTs from the MSU cDNA libraries were downregulated (37/6 236 = 0.6 %), but only nine ESTs from the UA libraries were strongly downregulated (9/2 976 = 0.3 %). Up- or downregulation in this experiment was defined as having a > 3-fold change in signal

Table III. Abundance profile of ESTs in a leaf cDNA library of salt-stressed rice cv. Pokkali (OG).

Gene category	No. of ESTs	% of total	Category description
No hit	16	3.4	BLAST search in non-redundant GenBank finds no homologs
Unknown	97	20.8	BLAST search GenBank identifies unknown, hypothetical, and putative proteins, or proteins with unknown functions
Metabolism	37	7.9	Amino acid, nucleotide, C-compound and carbohydrate, lipid, fatty acid and isoprenoid, nitrogen and sulfur, and secondary metabolisms
Cell rescue, defense	29	6.2	Environmental stimuli responses (stress, wounding, phytohormone regulation, and ion deficiency), radical scavenging, detoxification, DNA repair, and cell death
Protein synthesis	27	5.8	Ribosomal proteins, translation (initiation, elongation and termination), translational control, and tRNA synthetases
Energy	13	2.8	Glycolysis and gluconeogenesis, pentose-phosphate pathway, TCA pathway, respiration, metabolism of energy reserves (e.g. glycogen, trehalose), and fatty acid oxidation
Transport facilitation	11	2.4	Ion channels, water channels, ion, sugar and carbohydrate, amino acid, lipid, purine and pyrimidine transporters, transport ATPases
Protein destination	91	20.4	Protein folding and stabilization, protein targeting, sorting and translocation, protein modification (myristylation, farnesylation, palmityolation, glycosylation) and proteolysis
Signal transduction	21	4.5	Receptor proteins, second messenger such as calmodulins, key kinases, key phosphatases, and G-proteins
Cell growth, division	10	2.1	Cell growth, development, cell cycle control and mitosis, cytokinesis, and DNA synthesis, replication, and repair
Transcription	13	2.8	mRNA synthesis (including general transcription activities, transcriptional control such as transcriptional factors, and chromatin modification), mRNA splicing, and mRNA stabilization and degradation
Photosynthesis	156	33.5	Chlorophyll *a*/*b*-binding protein, ferredoxin related genes, photosystem I and II reaction center
Metal ion homeostasis	0	0	Homeostasis of metal ions and other ions
Cell structure	16	3.4	Cytoskeletal and other structural proteins (e.g. actin, histones)
Transposons	5	1.1	Transposition-related activities (En/Spm-like proteins)
Total	446	100	

intensity, in at least three out of four spots on two replicate hybridizations. Treatment was 150 mM NaCl for 24 h on 30-d-old plants grown on a slanted mesh.

2.3. *Arabidopsis* mutants in stress-relevant pathways

Generating mutant lines with a loss or gain of function constitutes the most parsimonious strategy for understanding plant responses to stress. We have generated a large number of tagged mutants of *Arabidopsis* using both simple T-DNA insertion vectors and activation vectors deployed with activation promoters comprised of enhancer sequences from both 35S promoter [4, 48] and super-promoter [35] origins. These insertions have been made using floral transformation with Columbia, C-24, and WS ecotypes. Over 400 000 independent insertion lines have been isolated and organized into ten line pools for potential reverse genetic screens. In addition, many of these insertion mutants have been made using plants with a stress responsive promoter (RD29A) fused to the marker gene LUC (luciferase) [19]. Altered luciferase expression compared to wild type plants reports mutations in a pathway affecting the environmental responsive characteristics of the RD29A promoter. We have begun to screen these mutants for various phenotypes. Salt tolerant and sensitive mutants are screened using the root bending assay described by Shi et al. [42]. Screening for salt tolerance is also accomplished by examining germination and growth in the presence of high NaCl.

Three categories of mutants have been identified using the RD29A promoter; constitutive osmotic sensitivity (COS), high osmotic sensitivity (hos), and low osmotic sensitivity (los), where the responsiveness of the promoter has either been increased (hos), decreased (los), or no longer requires induction (COS). HOS/LOS mutants are screened using an ultra-sensitive camera (custom-built “Back-illuminated” CCD camera, Roper Scientific) to detect variation in light

Table IV. Stress Model Organism EST PipeOnline_v2.0b database descriptions (based on the assumption that libraries are mRNA populations representative and tags selected through random sampling). [a] POL_db and the PipeOnline_v2.0b automated DNA sequence processing and functional sorting package are described (stress-genomics.org) (Ayoubi P., Jin X., Leite S., Liu X., Martajaja J., Abduraham R., Wan Q., Yan W., Misawa E., Prade R.A., unpubl.). Full and unrestricted database access for all the models listed are available from the stress-genomics.org website. [b] ESTs downloaded from dEST.

Surveyed EST libraries	ESTs		POLdb record with HSP > 100[a]					
	All	Unique	Total		Function		No function	
			No.	%	No.	%	No.	%
Plant models								
Arabidopsis thaliana	3 336	2 379	1 495	63	371	25	1 124	75
Hordeum vulgare	576	537	308	57	95	31	213	69
Mesembryanthemum crystallinum	7 327	3 788	2 096	55	567	27	1 529	73
Oryza sativa	3 816	2 553	1 224	48	322	26	902	74
Dunaliella salina	2 052	1 243	622	50	201	32	421	68
Selaginella sp.	1 191	995	769	77	215	28	554	72
All plant models	18 298	11 495	6 514		1 771		4 743	
Plant model average	3 050	1 916	1 086	58	295	28	791	72
Genomic model systems								
Aspergillus nidulans[b]	12 485	4 595	1 863	41	609	33	1 254	67
Saccharomyces cerevisiae[b]	2 799	1 587	1 343	85	318	24	1 025	76
All genomic models	15 284	6 182	3 206		927		2 279	
Genomic model average	7 642	3 091		63		28		72
All biological systems	30 783	16 090	8 377		2 380		5 997	
Overall average	3 848	2 011		61		28		72

produced by the luciferase marker. *Table VI* summarizes the number of mutants in the *hos*, *los* and *sos* categories that have been isolated and their frequencies of appearance. Shown also (*table VII*) are the frequencies of other morphological or developmental mutants that have been isolated from the insertion tagged populations. Although the tagged populations have not been screened exhaustively for types of visible mutants, these frequencies can indicate a minimum frequency of phenotypic mutations occurring in the population. Making the T-DNA tagged lines available to the community in the form of DNA super-pools and as populations of multiplied seeds is a major objective of our work.

2.4. Analysis of non-plant models

One part of our work on global salt stress-related gene expression changes is the comparative analysis of transcript expression in non-plant model species. We expect a phylogenetic cross section of stress-regulated genes or ORFs to provide an appreciation of the mechanisms that need to be engaged when plants experience high salinity. This sectioning is possible because the genomes of several of the models used by us are available. The genomes of *Synechocystis* PCC6803 and *Saccharomyces cerevisiae* are completely sequenced, and those of *Aspergillus nidulans*, *A. thaliana* and *Oryza sativa* are available to some degree, with that of *Arabidopsis* to more than 90 % accessible. *Table VIII* is anchored on ORFs in yeast that are induced by salinity stress [50] in an experiment that compared yeast cells stressed by 1 M NaCl for 10, 30 and 90 min. The table shows yeast ORFs that are upregulated after 10 min stress and those transcripts that are also represented in cDNA libraries from salt-stressed *Arabidopsis*, rice and the ice plant.

Outstanding in this list of yeast ORFs and their plant orthologs is the large number of ribosomal proteins (RPO) that are induced in yeast within 10 and 30 min after stress [50]. Later during stress in yeast, these *rpo*-transcripts are downregulated. A very similar observation has been made in salt-stressed rice (Kawasaki S., Deyholos M., Borchert C., Brazille S., Kawai K., Galbraith D.W., Bohnert H.J., unpubl.) with the exception that upregulation of the *rpo*-transcripts was strongest 1 and 3 h after the beginning of the stress, while downregulation was again observed later. A comparison between rice and yeast seems to indicate that the synthesis of ribosomes and, coincident with this synthesis, the elevation of transcripts for protein turnover and re-structuring of the proteome are early events following salt shock. Ubiquitously present in yeast and

Table V. Selected ESTs highly expressed in either EBC or shoots in *Mesembryanthemum* (data from microarray slides containing ~2 600 EST elements). Ratio shoot/EBC and EBC/shoot for a selected number of transcripts from two independent hybridization experiments are shown. EBC, epidermal bladder cells.

Expt 1	Expt 2	Clone ID#	Annotation	Homology	Functional categories
EBC-specific ESTs					
12.8	34.4	MH03A05	anti-fungal protein 1 precursor	P81418 pokeweed	Cell rescue, defense, cell death, and aging
9.4	12.3	ML02G10	annotation in progress	–	NA
6.9	8.8	MO01F11	carboxyphosphonoenolpyruvate mutase	S35145 carnation	Metabolism
9.7	8.5	MG04B05	major allergen Mal d1, pathogenisis-related	S11869 potato	Cell rescue, defense, cell death, and aging
29.3	8.1	MH02C01	antimicrobial peptide, AMP2	P81418 pokeweed	Cell rescue, defense, cell death, and aging
2.5	7.4	MH01C03	pathogenesis-related protein 5 precursor, thaumatin	P28493 *Arabidopsis*	Cell rescue, defense, cell death, and aging
9.3	6.9	MM04F08	major allergen Mal d1, pathogenisis-related	Z72425 apple tree	Cell rescue, defense, cell death, and aging
3.8	6.5	MH02C02	receptor-like protein kinase	AC000132 *Arabidopsis*	Signal transduction
4.9	5.3	MF07F05	no hits	–	NA
2.1	4.8	MP09D06	no hits	–	NA
4.0	4.8	ML01E05	no hits	–	NA
4.9	4.6	ML03C02	non-specific lipid-transfer protein	Q43748 sugarbeet	Cellular organization
8.8	4.6	MK01E03	no hits	–	NA
6.9	4.1	MM01G02	unknown	AB003280 *Arabidopsis*	NA
8.6	3.5	MM02F06	annotation in progress	–	NA
Shoot-specific cDNAs					
24.6	3.9	MP03C10	unknown, desiccation-related protein 2	P22242 *C. plantagineum*	Cell rescue, defense, cell death, and aging
15.2	3.8	MH02D01	BURP-domain containing protein	AC007504 *Arabidopsis*	NA
16.4	3.9	MO03G01	vegetative storage protein, wound-induced	S39502 poplar	NA
9.8	3.8	MO02E11	lipid transfer protein, glossy (gl1) homolog	L33792 *S. odorus*	Cellular organization
8.9	3.8	MP07B01	water channel, tonoplast membrane, MIP-F	U43291 ice plant	Transport facilitation
7.3	3.8	MG01A01	no hits	–	NA
6.9	3.8	ME02E06	unknown, jasmonate-induced protein	P42764 *Atriplex*	NA
6.3	3.7	MG04B01	no hits	–	NA
5.7	3.7	MG01A07	unknown	AF024504 *Arabidopsis*	NA
5.2	3.7	MP14F03	water channel, plasma membrane, MIP A	L36095 ice plant	Transport facilitation
4.9	3.6	R30-tip#2	water channel, tonoplast membrane, MIP I	AF133531 ice plant	Transport facilitation
4.9	3.6	R30-tip#18	lipoxygenase	AF145479 ice plant	Metabolism/signal transduction
4.8	3.6	MP14B04	chitinase	JC4053 pokeweed	Cell rescue, defense, death, and aging
4.8	3.6	MP01G05	cysteine proteinase	U30322 ice plant	Protein destination
4.8	3.6	MP14C02	unknown	P13729 fruit fly	NA

the plant models are transcripts for functions in the categories of cell defense and cell rescue and these transcripts are upregulated in all species studied by us. The time course of upregulation seems to play a major role, however. For example, in rice the very early upregulated functions show homology with protein kinases and other putative signal transduction intermediates, followed by functions that seem to re-structure protein synthesis and protein turnover. Later, the well-known functions associated with elevated ABA appear among the upregulated ESTs in rice. ESTs with putative functions in radical oxygen scavenging are induced at several time points following stress (Kawasaki S., Deyholos M., Borchert C., Brazille S., Kawai K., Galbraith D.W., Bohnert H.J., unpubl.). A succession of upregulation of different functions is reminiscent of the phases of stress responses, labeled immediate turgor-related, followed by metabolic adjustments and later developmental changes, that have been outlined before [34]. The experiments so far have indicated a progression of upregulated ESTs in different functional categories, indicating that a reliance on single time points (and single hybridizations) will not provide sufficient information on the complexity of

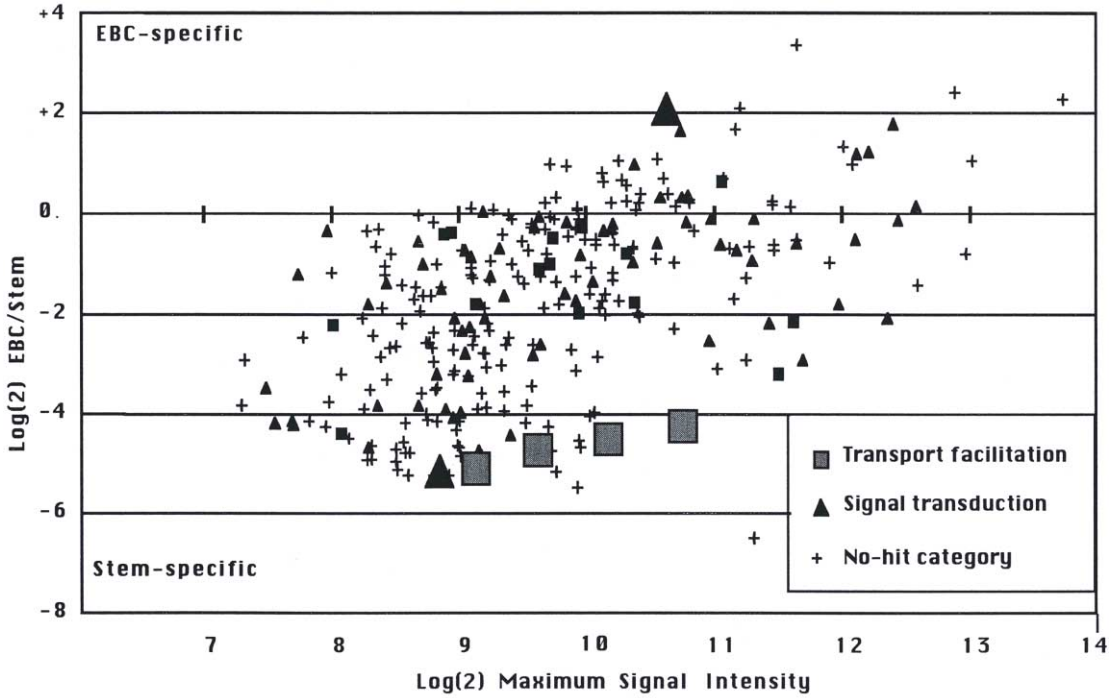

Figure 2. Comparison of EST expression in different tissues. Selected ESTs from a total of 2 600 ESTs from salt-stressed ice plants are shown in three categories: (+), no hit; (triangle), transport facilitation; (square), signal transduction. The log(2) ratio of transcript abundance in epidermal bladder cells (EBC) vs. stems is shown vs. log(2) of maximum signal intensity. Most highly expressed ESTs tend to group in EBC (top right), stem-specific transcripts are mostly of lower intensity (bottom left). Four highlighted, stem-specific ESTs show high similarity to known ice plant water channel transcripts. The two enlarged triangles show closely related ESTs putatively encoding for MAP-kinases which are distinguished by the array hybridizations (according to Kawasaki S., Deyholos M., Borchert C., Brazille S., Kawai K., Galbraith D.W., Bohnert H.J., unpubl.).

gene expression changes under osmotic and ionic stress conditions.

Table IX presents a set of *Synechocystis* sp. PCC6803 genes [24] with significant homologies (expectation values < 0.1) to a set of ESTs obtained from salt-stressed higher plant and other model organism cDNA libraries. The table represents a subset of the overall BLASTX search that obtained 1 964 'hits' meeting this criterion of homology out of a 11 194 unigene set of ESTs. Clearly, several expected categories of proteins such as those having stress and transport functions are well represented in the *Synechocystis* PCC6803 genome. These ORFs are being targeted for knock-out and stress phenotype analysis [6, 49]. One important aspect of the analysis that is not conveyed by this table is the large number of genes falling into the category of 'hypothetical' and 'putative' proteins or proteins of unknown function. In fact, over 400 (20 %) of the higher organism stress-related ESTs originating from this search are listed as hypothetical or unknown proteins. Therefore, in addition to better defining the function of known genes in mitigating salt stress in oxygenic photosynthesis, *Synechocystis* sp. PCC6803 will be useful to define the functions of unknown coding regions [6, 49]. The analysis of gene function

Table VI. Frequencies of mutants isolated as variants in responsiveness of the RD29A promoter or as salt tolerant or sensitive. All indicated mutants have been confirmed by a second screen of progeny from self-fertilized T_2 mutants.

Phenotype of mutant	Population screened
hos/los	80 000
43 hos	
20 los	
sos3 suppressors	55 000
5 *sos3* suppressors	
13 suppressors of part of the *sos3* syndrome	
Salt tolerant	10 000
3	
Salt sensitive	50 000
12	

Table VII. Screens of 35 000 (T_2) T-DNA insertion mutant lines of C24 ecotype. All mutants were generated by transformation with the pSKI vector described by Weigel et al. [48] and confirmed in progeny of T_2 plants except as indicated. Many possible visible phenotypes were not carefully examined. For example, populations were not inspected for floral structure variants. Agamous-like variants were observed as continuous flowering.

Phenotype	No. found in T_2 pop.
Dwarf	11
Eceriferum	7
Glabrous	9
Elongated spatulate leaves	3
Thin strap-like leaves and petals	2
Wrinkled/cupped leaves	7
Cauliflower inflorescence	1
Agamous-like	2
Late flower	9
Pin-like	2
Dark green leaves	5
Light green leaves	3
Small leaves	4
Round leaves	2
Multi-branch shoot	5
Curled leaf	2
Upright siliques	1
Reduced fecundity	20
Miscellaneous unverified/steriles	33
Putative mutants that were wild type in T_3	58

can in many cases, as for example putative transport and membrane proteins with homology across phyla, be accomplished by engineering approaches and complementation using the *Synechocystis*, *Aspergillus* and yeast models.

2.5. Towards functional analysis

Approximately half of all ESTs, cDNAs and genes that have become available are functionally not characterized, and the frequency of functionally unknown transcripts in cDNA libraries from stressed tissues is even higher. Strategies that address function for those transcript sequences are still emerging and the existing approaches target different objectives. An initial goal is knowledge about the organs, tissues and cells in which the functionally unknown genes are expressed, and the time of their expression during development. Complementing microarray analysis by using high-throughput transformation of promoter-reporter gene constructs in Arabidopsis and rice will help to achieve a global blueprint of promoter activity (e.g. [22, 30]). Functional complementation in model organisms such as *Synechocystis*, yeast, *Aspergillus* and *Arabidopsis* mutants presents a second strategy. Large-scale functional complementation assays of mutant phenotypes, correlated with microarrays or DNA-chips that include the entire transcriptome in comparisons between wild type, mutant and complemented mutant, are being designed. This goal is not yet within reach. The available microarrays and DNA-chips provided by companies, by a group at Michigan State University (Green P., pers. comm.) and the microarray slides which we have assembled contain only approximately one-third of the transcripts that are assumed to constitute the Arabidopsis transcriptome. With a population of approximately 400 000 T-DNA tagged mutants from Arabidopsis, saturation mutagenesis of this genome has most likely been achieved but the task of organizing and providing such a large number of mutant lines requires time.

3. CONCLUSION

With the instrumentation in place that allows large-scale DNA sequencing, genome-wide analysis of the transcriptome, and large-scale analysis of all proteins in a cell, we must learn to utilize the bioinformatics tools that are essential if we wish to sort and understand the data avalanche. The extent to which biology is presently driven by these new technologies and by the need for sophisticated bioinformatics tools requires a new mindset. We have arrived at a threshold that permits the analysis of plant stress responses on the level of entire genomes. ESTs, microarrays and the clustering of expression patterns, the availability of molecular markers and mutants, and the mapping of robust quantitative trait loci (QTLs) provide tools for assessing the contribution of many genes to stress responses and tolerance under stress [8, 12, 14, 25, 26, 38]. In addition, tools for the global analysis of metabolites have emerged [46, 47]. Conclusions can now be drawn that supersede correlative evidence. A major effort will be necessary to integrate the diverse elements of information into comprehensive databases. The use of cluster analysis (e.g. [8]) that allows for integration of the expression characteristics of thousands of transcripts is especially expected to provide insight. Clusters of similarly regulated transcripts, which include functionally known and unknown transcripts, will help in the formation of hypotheses that eventually will permit functional identifications of unknown coding regions. Understanding plant stress tolerance on the level of genome-wide responses is several years in the future. The first candidate genes that could be essential components supporting abiotic

Table VIII. Plant ESTs homologous to salt stress-induced yeast ORFs. *Yeast ORFs upregulated during salt stress (1 M; 10, 30 and 90 min) are described [50]. **For MIPS classification see http://websvr.mips.biochem.mpg.de/proj/yeast/catalogues/funcat/.

Systematic name	Gene name	Functional annotation	MIPS category**	*Mesembryanthemum*	Rice	Corn
YPL218W	SAR1	GTP-binding protein of the ARF family	INTTRA	ME01G01		
YJR009C	TDH2	glyceraldehyde-3-phosphate dehydrogenase 2	MET	ME02A12	OD105G05	
YAL004W*		strong similarity to *A. klebsiana* glutamate dehydrogenase	U	ME02D10		
YAL005C*	SSA1	heat shock protein of HSP70 family, cytosolic	CDEF	ME02D10	OC11H04	
YKR014C	YPT52	GTP-binding protein of the *rab* family	PDEST	ME02H03	OC104H06	
YBL087C	RPL23A	60S large subunit ribosomal protein L23.e	PS	ME03G02	OC10A11	ZA5740407
YER103W	SSA4	heat shock protein of HSP70 family, cytosolic	CDEF	ME04D06		
YGL166W*	CUP2	copper-dependent transcription factor	TRANSC	MF02C06		
YJL190C	RPS24A	ribosomal protein S15a.e.c10	PS	MF02C10	OC01A08	
YHR049C-A*		questionable ORF	U	MF02E11		
YPL079W	RPL21B	ribosomal protein L21	PS	MF03B08		
YEL059W		hypothetical protein	U	MF04B05		
YIL052C	RPL34B	ribosomal protein L34.e	PS	MF05B01		
YNL192W	CHS1	chitin synthase I	MET	MF06B02		
YER130C		similarity to MSN2P and weak similarity to MSN4P	U	MF07E01		
YGR137W		questionable ORF	U	MF07G10		
YMR110C		similarity to aldehyde dehydrogenase	U	MG01A11		
YDR055W	PST1	strong similarity to SPS2 protein	U	MG01D06		
YLR167W	UBI3	ubiquitin/40S small subunit ribosomal protein	PS	MG01D09	OE11C03	ZA0C265
YER141W	COX15	cytochrome oxidase assembly factor	MET	MG01E08		
YPR035W	GLN1	glutamate-ammonia ligase	MET	MG01E12		
YJR121W	ATP2	F1F0-ATPase complex, F1 beta subunit	IONHOM	MG02C10	OC03G05	
YDR064W	RPS13C	ribosomal protein	PS	MG04E09		
YDR461W	MFA1	mating pheromone a-factor 1	CGRO	MG05D01		
YGL165C*		questionable ORF	U	MG05D03	OF03D10	
YHR137W*	ARO9	aromatic amino acid aminotransferase II	MET	MG05F04		
YMR186W	HSC82	heat shock protein	CDEF	MH01B05	OC09E02	
YLR029C	RPL15A	RPL15A 60s large subunit ribosomal protein L15.e.c12	PS	MH01G05		
YKL036C*		questionable ORF	U	MH03B02		
YLR333C	RPS31	ribosomal protein S25.e.c12	PS	MH04B07		
YPL090C	RPS6A	ribosomal protein S6.e	PS	MH05B07	OE08E07	
YOL087C		similarity to *S. pombe* hypothetical protein	U	MH05E12		
YLR293C	GSP1	GTP-binding protein of the *ras* superfamily	TRANSC	MH05H04		ZA4730320
YAL003W	EFB1	translation elongation factor eEF1beta	TL		OC01C02	
YGL248W	PDE1	low affinity 3',5'-cyclic-nucleotide phosphodiesterase	MET		OC02B01	
YGR085C	RPL11B	ribosomal protein	PS		OC08H01	
YLL026W	HSP104	heat shock protein	CDEF		OC101B09	
YML001W	YPT7	GTP-binding protein of the RAB family	INTTRA		OC104H06	
YPL081W	RPS9A	ribosomal protein S9.e.A	PS		OC105A09	ZA5740676
YEL054C	RPL15B	60S large subunit ribosomal protein L12.e	PS		OC105B04	
YGL147C	RPL9A	ribosomal protein L9.e	PS		OC10B09	ZA5739836
YKL100C		similarity to *C. elegans* hypothetical protein	U		OD103H07	
YHR203C	RPS7A	RPS4B ribosomal protein S4.e.c8	PS		OE03C03	
YDR154C*		questionable ORF	U		OE05B12	
YDR155C*	CPH1	cyclophilin (peptidylprolyl isomerase)	PDEST		OE05B12	
YDR025W	RPS18A	ribosomal protein S11.e	PS		OE07F09	
YDR304C	CYP5	peptidyl-prolyl cis-trans isomerase D precursor(cyclophilin D) of the ER	MET		OE10D02	
YOL040C	RPS21	RPS15, 40S small subunit ribosomal protein	PS		OE14G07	
YMR119W-A		Questionable ORF	U		OE202E11	
YER067W		Strong similarity to hypothetical protein YIL057c	U		OF03D09	
YJL053W	PEP8	Vacuolar protein sorting/targeting protein	PDEST			ZA0C131
YJL138C	TIF2	Translation initiation factor eIF4A	PS			ZA0C157
YEL024W	RIP1	Ubiquinol-cytochrome-c reductase iron-sulfur protein precursor	ENERGY			ZA0C93
YBL027W	RPL19A	60S large subunit ribosomal protein L19.e	PS			ZA5739792
YGR086C		strong similarity to hypothetical protein YPL004c	U			ZA5757355

Table IX. Selected *Synechocystis* PCC6803 genes with homology to yeast stress-induced ORFs and higher plant stress-related ESTs. Table representing a selected list of BLAST hits from higher plant salt-stressed EST sequences against the *Synechocystis* PCC6803 (whole genomic) amino acid sequence. Of the 11 194 salt stress EST's queried, 1 964 show homology to *Synechocystis.* Tissue types are as follows: R, roots; L, leaves; F&S, flowers and seedpods; S, seedlings; BC, bladder cells; WS, whole seedlings; M, meristems.

Functional category	*Synechocystis* gene	Yeast	Rice	Barley	Ice plant	Arabidopsis
Light regulated	CAB/ELIP/HLIP superfamily	–	–	–	R	L
	highlight-inducible protein	–	–	–	'M, F&S'	L
Chaperonins	DnaJ protein	–	R	'L, R'	'M, WS'	L
	plastocyanin	–	–	–	'M, F&S'	'L, R, WS'
	10-kD chaperonin	–	–	–	'M, BC'	L
	DnaK protein	Y	R	R	All	'L, R'
	60-kD chaperonin	–	–	–	M	All
Oxidative stress	catalase hpi (katG)	Y	–	R	'R, F&S, M'	'L, R'
	thiol-specific antioxidant protein	–	R	–	M	'L, R'
	glutaredoxin 3	Y	R	–	'R, M '	L
	glutathione peroxidase	Y	–	L	2' L	'L, R'
	membrane protein	–	R	–	R	L
	superoxide dismutase	Y	–	–	M	L
	Ycf39	–	–	–	R	'L, R, WS'
Transporters	Na^+/H^+-antiporter	Y	–	–	–	–
	Na-ATPase (PacL)	Y	–	–	–	–
	ABC transporter	–	R	–	'R, M, F&S'	–
	water channel	–	R	R	'R, M, S'	All
	spore protein sp21	–	–	R	'R, M'	–
	salt-stress hydrophobic peptide	–	R	–	'R, M, BC'	'L, R, WS'
	high affinity sulfate transporter	Y	–	–	–	WS
	cation-transporting ATPase PacL	–	–	–	BC	L
Translational factors	initiation factor IF-2	–	R	R	'R, M, F&S'	'L, R'
	RNA-binding protein	–	R	L	'R, M, F&S, BC'	all
	elongation factor Tu	–	R	'L, R'	'R, M, F&S'	L
Kinases and phosphatases	monophosphatase	Y	–	–	–	–
	extragenic suppressor SuhB	Y	R	–	–	'L, R'
	protein kinase PknA	Y	–	–	R	R
	protein kinase C inhibitor	–	R	–	R	–
	serine esterase	Y	–	–	R	–
	eukaryotic protein kinase	Y	R	R	'R, M, F&S'	L
	sensory transduction histidine kinase	–	R	–	–	'L, R'
Carbon concentration	betatransducin-like protein	–	R	–	'R, M'	'L, R, WS'
Energy metabolism	ferredoxin-NADP oxidoreductase	Y	R	–	–	'L, WS'
	UDP-glucose dehydrogenase	–	R	'L, R'	–	'L, R'
	S-adenosylhomocysteine hydrolase	–	R	R	'R, M, F&S'	'L, R'
Hypothetical	Hypothetical proteins	Y	many	many	many	many

stress tolerance have appeared in a number of stress signal recognition and transduction components and in additional downstream reactions, such as functions in sodium partitioning [3, 15, 17, 19, 23, 33, 42, 43] but additional components can be expected.

4. METHODS

4.1. Libraries from stressed plants and EST sequencing

Our main objective was to obtain cDNA libraries from different tissues of salt-stressed plants to be used for the generation of EST sequences. Up to now, approximately fifty cDNA libraries have been generated, listed at stress-genomics.org (see also *table I*). They include cDNA libraries from the non-plant models *Synechocystis* sp. PCC6803, *Saccharomyces cerevisiae*, *Aspergillus nidulans* and *Dunaliella salina*. The plant models for which cDNA libraries have been established are *Arabidopsis thaliana*, *Hordeum vulgare*, *Mesembryanthemum crystallinum*, *Oryza sativa* and *Zea mays*.

For cDNA synthesis, $polyA^+$ RNA was isolated using Dynabeads Oligo $(dT)_{25}$ (DYNAL, Inc.) in three

binding- and release-cycles. Libraries were constructed from 2.5 µg RNA by reverse transcription (cDNA Synthesis Kit; Stratagene), ligated into pBluescript II XR (Stratagene) for directed cloning and plated onto Xgal/IPTG/Amp plates for blue/white selection.

4.2. Template preparation, DNA sequencing and data analysis

Plasmid templates were prepared from selected bacterial colonies by 96-well alkaline lysis minipreps according to the manufacturer's instructions (Edge BioSystems, Inc. Gaithersburg, MD). DNA sequencing reactions were conducted using DyeDeoxy™ Terminator PRISM™ mix (Perkin-Elmer-ABI, Inc. Foster City, CA) according to the manufacturer's instructions in Multiplate™ thin-wall 96-well microplate on a MJ Research PTC-100-96 (MJ Research, Inc., Watertown, MA) programmable thermal controller (96 °C for 30 s, 45 °C for 15 s, 60 °C for 4 min for 49 cycles). Unincorporated dye terminators were removed by filtration (96-well filtration block, Edge BioSystems, Inc.). Reaction products were sequenced on either the ABI 373A-XL Stretch or ABI 3700 capillary array automated DNA sequencing systems (Perkin-Elmer Applied BioSystems, Inc.). Raw sequence data was analyzed using PHRED [10, 11] and cross-match to remove vector sequences. Additional sequence data polishing was done manually using FACTURA™ software (Perkin-Elmer Applied BioSystems, Inc.). Polished ESTs sequence files were assembled into singleton and contig files using PHRAP (Green P., unpubl.). EST identities were determined by sequence comparison to the non-redundant GenBank database using BLASTN (BLAST 2.0) using default parameters [2]. Stress-related EST sequences that have been deposited in dbEST may be retrieved from www.stress-genomics.org.

4.3. Microarray generation, hybridization and analysis

4.3.1. Probe preparation

Inserts were amplified from EST libraries in 96-well format, using vector-specific primers (SigmaGenosys, Woodlands, TX), in standard 50 µL PCR reactions. The EST libraries came from our laboratories and (for Arabidopsis) from the Arabidopsis Biological Resource Center (Columbus, OH). We use 96-well MAF-NOB™ glass-fiber filter plates (Millipore, Bedford, MA) to purify probes from residual components of the PCR reaction, by first resuspending the PCR mixture in an equal volume of binding buffer (7 M Guanidine HCl, 200 mM MES free acid, 20 mM MES sodium salt, pH 5.6), then drawing the mixture through the filter plate on a vacuum manifold. The filter-bound DNA was washed twice with 80 % ethanol, dried, and then DNA was eluted from the filters with 120 µL 0.1× TE. An aliquot of each purified probe was analyzed by agarose gel electrophoresis, and the remainder was desiccated in a SpeedVac® (Savant Instruments, Holbrook, NY) and resuspended in 6 µL 2× SSC at a final DNA concentration of at least 300 $ng \cdot \mu L^{-1}$.

4.3.2. Probe printing and immobilization

Probes were spotted on either SigmaScreen™ or PolyPrep™ slides (Sigma Chemical, St Louis, MO) using an OmniGrid™ robot (GeneMachines, San Carlos, CA) equipped with quill-type steel pins (Majer Precision, Tempe, AZ; and Telchem, Sunnyvale, CA). For our highest-density arrays, spots were printed at a nominal center-to-center spacing of 180 µm. After spotting, arrays were stored at room temperature for up to 4 months before immobilization. Prior to hybridization, we immobilized and blocked the arrays as follows. Arrays were rehydrated by holding them approximately 2 cm above the surface of a 42 °C water bath for 3–5 s, and then were snap-dried by placing them face-up on a 65 °C heat block. The arrays were subsequently exposed to 65 $mJ \cdot cm^{-2}$ UV in a Stratalinker® (Stratagene, La Jolla, CA), then rinsed in 1 % SDS for 2 min, and denatured by placing them in boiling, deionized water for 2 min. Slides were transferred from the boiling water to 95 % ethanol for 1 min, then dried by centrifugation for 30 s at 2 000 × *g* in a swinging bucket centrifuge (Beckman, J6).

4.3.3. Target preparation

Total RNA (and occasionally $polyA^+$ RNA) was labeled by direct incorporation of Cy3- or Cy5-conjugated dUTP (Amersham Pharmacia Biotech, Piscataway, NJ) during reverse transcription as follows. Total RNA (150 µg) was combined with 0.5 µg anchored oligo $(dT)_{23}$, heated to 70 °C for 10 min and then cooled on ice. To this mixture, the remaining components were added to obtain the following reaction conditions in 30 µL total volume: 1× AMV reverse transcriptase buffer, 20 U Enhanced Avian Reverse Transcriptase (Sigma Chemical, St Louis, MO), 20 U RNase inhibitor (Sigma Chemical, St Louis, MO), 0.03 mM Cy3- or Cy5-dUTP, 0.33 mM each (dATP, dCTP, dGTP) and 0.07 mM TTP (Sigma Chemical, St Louis, MO). Following a 90-min incubation at 42 °C,

the reaction was stopped by adding 2.5 µL 0.5 M EDTA, and the RNA was hydrolyzed by adding 5 µL 1 M NaOH and incubating for 10 min at 70 °C. This reaction was neutralized by adding 12.5 µL 1 M Tris-HCl (pH 7.5). The labeled target was purified from unincorporated dye molecules by adding 400 µL TE, and spinning through a Microcon YM-30 filter (Millipore, Bedford, MA) for 15 min at 11 000 × *g*. The purified labeled target was concentrated to a final volume of less than 7.5 µL.

4.3.4. Hybridization and washing

For hybridization to the *Arabidopsis* array, Cy3- and Cy5-labeled targets were combined in 4 µL 20× SSC, 1 µL 2 % SDS and 2 µL Liquid Block (Amersham Pharmacia Biotech, Piscataway, NJ) to a final volume of 22 µL. The mixture was placed in a boiling water bath for 2 min, cooled on ice, then applied to the surface of a blocked and immobilized microarray that had been pre-warmed to 65 °C. The microarray and hybridization solution were overlaid with a 22 × 40 mm Hybri-Slip and transferred to a pre-warmed 50-mL screw-top disposable centrifuge tube that was kept humid by enclosing a folded paper towel moistened with 2× SSC. The array was incubated in 65 °C air for 12 h. Following hybridization, the microarray was washed at 65 °C in 2× SSC, 0.5 % SDS for 5 min, then at room temperature in 0.5× SSC for 5 min, and finally at room temperature in 0.05× SSC for 5 min. Slides were dried by spinning at low speed in a centrifuge for 30 s.

4.3.5. Scanning and analysis

The arrays were scanned in a ScanArray® 3000 (GSI Lumonics, Billerica, MA). Laser and PMT voltages were adjusted manually to minimize background and the number of spots that returned saturated signal values. The scanner output images were quantified using Imagene™ 3.04 (BioDiscovery, Los Angeles, CA). For standardization, the Cy3 and Cy5 signal intensities were adjusted relative to each other by comparing the mean signal intensity of all spots in each channel.

4.4. Mutants in stress-relevant genes

The generation of T-DNA tagged mutants has been described [20, 35,48].

Acknowledgments. We thank the technical staff and students in our laboratories, who have been essential in making high-throughput DNA preparation, analysis of ESTs, and mutant generation possible, especially Rashid Abduraham, Inna Akselrod, Adam Barb, Alison Berne, Susan Brazille, Amanda Breen, Davi Bufford, Juang-Hong Chong, Michael Dellinger, Monica Dennis, Andrea Hanlon, Xiaojing Jin, Melissa Knue, Julianne Lazarr, Saul Leite, Fang Li, Xianghui Liu, Heather McLaughlin, Jeson Martajaja, Jessica Martinez, Michael McCully, Jill Montgomery, Emmanuel Mulenga, Chris Palacio, George Risinger, Janet Rogers, Stacia Sant, Irina Sokolchik, Briana Sollod, Tema Starkman, JoAnne Thomas, Qiaolan Wan, Tara Ware, Chieh-Kuen Woo, Wei Yan. The work is supported by the US National Science Foundation, Plant Genome Program (award 98-13360).

REFERENCES

[1] Adams P., Nelson D.E., Yamada S., Chmara W., Jensen R.G., Bohnert H.J., Griffiths H., Growth and development of *Mesembryanthemum crystallinum* (Aizoaceae), New Phytol. 138 (1998) 171–190.

[2] Altschul S.F., Madden T.L., Schäffer A.A., Zhang J., Zhang Z., Miller W., Lipman D.J., Gapped BLAST and PSI-BLAST: a new generation of protein database search programs, Nucleic Acids Res. 25 (1997) 3389–3402.

[3] Apse M.P., Aharon G.S., Snedden W.A., Blumwald E., Salt tolerance conferred by overexpression of a vacuolar Na^+/H^+ antiport in *Arabidopsis*, Science 285 (1999) 1256–1258.

[4] Barakat A., Gallois P., Raynal M., Mestre-Ortega D., Sallaud C., Guiderdoni E., Delseny M., Bernardi G., The distribution of T-DNA in the genomes of transgenic *Arabidopsis* and rice, FEBS Lett. 471 (2000) 161–164.

[5] Bohnert H.J., Su H., Shen B., Molecular mechanisms of salinity tolerance, in: Shinozaki K. (Ed.), Cold, Drought, Heat, and Salt Stress: Molecular Responses in Higher Plants, R.G. Landes, Austin, 1999, pp. 29–60.

[6] Burnap R.L., Troyan T., Sherman L.A., The highly abundant chlorophyll-protein complex of iron-deficient *Synechococcus* sp. PCC7942 (CP43') is encoded by the *isiA* gene, Plant Physiol. 103 (1993) 893–902.

[7] Cushman J.C., Bohnert H.J., Genomic approaches to plant stress tolerance, Curr. Opin. Plant Biol. 3 (2000) 117–124.

[8] Eisen M.B., Spellman P.T., Brown P.O., Botstein D., Cluster analysis and display of genome-wide expression patterns, Proc. Natl. Acad. Sci. USA 95 (1998) 14863–14868.

[9] Evans L.T., Feeding the Ten Billion, Cambridge University Press, 1998.

[10] Ewing B., Green P., Base-calling of automated sequencer traces using Phred. II. Error probabilities, Genome Res. 8 (1998) 186–194.
[11] Ewing B., Hillier L., Wendl M.C., Green P., Base-calling of automated sequencer traces using Phred. I. accuracy assessment, Genome Res. 8 (1998) 175–185.
[12] Ferea T.L., Botstein D., Brown P.O., Rosenzweig R.F., Systematic changes in gene expression patterns following adaptive evolution in yeast, Proc. Natl. Acad. Sci. USA 96 (1999) 9721–9726.
[13] Flowers T.J., Yeo A.R., Breeding for salinity tolerance in crop plants: Where next?, Aust. J. Plant Physiol. 22 (1995) 875–884.
[14] Frova C., Caffulli A., Pallavera E., Mapping quantitative trait loci for tolerance to abiotic stresses in maize, J. Exp. Zool. 282 (1999) 164–170.
[15] Halfter U., Ishitani M., Zhu J.K., The *Arabidopsis* SOS2 protein kinase physically interacts with and is activated by the calcium-binding protein SOS3, Proc. Natl. Acad. Sci. USA 97 (2000) 3735–3740.
[16] Hare P.D., Cress W.A., van Staden J., The involvement of cytokinins in plant responses to environmental stress, Plant Growth Regul. 23 (1997) 79–103.
[17] Hasegawa P.M., Bressan R.A., Zhu J.K., Bohnert H.J., Plant cellular and molecular responses to high salinity, Annu. Rev. Plant Physiol. Plant Mol. Biol. 51 (2000) 463–499.
[18] Horton P., Ruban A.V., Walters R.G., Regulation of light harvesting in green plants, Annu. Rev. Plant Physiol. Plant Mol. Biol. 47 (1996) 655–684.
[19] Ishitani M., Liu J., Halfter U., Kim C.S., Shi W., Zhu J.K., SOS3 function on plant salt stress tolerance requires N-myristoylation and calcium binding, Plant Cell 12 (2000) 1667–1678.
[20] Ishitani M., Xiong L., Stevenson B., Zhu J.K., Genetic analysis of osmotic and cold stress signal transduction in *Arabidopsis*: interactions and convergence of abscisic acid-dependent and abscisic acid-independent pathways, Plant Cell 9 (1997) 1935–1949.
[21] Jain R.K., Selvaraj G., Molecular genetic improvement of salt tolerance in plants, Biotechnol. Annu. Rev. 3 (1997) 245–267.
[22] Jeon J.S., Lee S., Jung K.H., Jun S.H., Jeong D.H., Lee J., Kim C., Jang S., Yang K., Nam J., An K., Han M.J., Sung R.J., Choi H.S., Yu J.H., Choi J.H., Cho S.Y., Cha S.S., Kim S.I., An G., T-DMA insertional mutagenesis for functional genomics in rice, Plant J. 22 (2000) 561–570.
[23] Liu J., Ishitani M., Halfter U., Kim C.S., Zhu J.K., The *Arabidopsis thaliana* SOS2 gene encodes a protein kinase that is required for salt tolerance, Proc. Natl. Acad. Sci. USA 97 (2000) 3730–3734.
[24] Kaneko T., Tabata S., Complete genome structure of the unicellular cyanobacterium *Synechocystis* sp. PCC6803, Plant Cell Physiol. 38 (1997) 1171–1176.
[25] Kehoe D.M., Villand P., Somerville S., DNA microarrays for studies of higher plants and other photosynthetic organisms, Trends Plant Sci. 4 (1999) 38–41.
[26] Lemieux B., Aharoni A., Schena M., Overview of DNA technology, Mol. Breeding 4 (1998) 277–289.
[27] Levitt J., Responses of Plant to Environmental Stress Chilling, Freezing, and High Temperature Stresses, 2nd ed, Academic Press, New York, 1980.
[28] Lin X., Kaul S., Rounsley S., Shea T.P., Benito M.I., Town C.D., et al., Sequence and analysis of chromosome 2 of the plant *Arabidopsis thaliana*, Nature 402 (1999) 761–769.
[29] Long S.P., Humphries S., Falkowski P.G., Photoinhibition of photosynthesis in nature, Annu. Rev. Plant Physiol. Plant Mol. Biol. 45 (1994) 633–663.
[30] Macas J., Lambert G.M., Dolezel D., Galbraith D.W., NEST Analysis: a novel means to study transcription through amplification of nuclear RNA, Cytometry 33 (1998) 460–468.
[31] McCullum C.M., Comai L., Greene E.A., Henikoff S., Targeting induced local lesions IN genomes (TILLING) for plant functional genomics, Plant Physiol. 123 (2000) 439–442.
[32] Miflin B., Crop improvement in the 21st century, J. Exp. Bot. 51 (2000) 1–8.
[33] Mizoguchi T., Ichimura K., Yoshida R., Shinozaki K., MAP kinase cascades in *Arabidopsis*: their roles in stress and hormone responses, Results Probl. Cell Differ. 27 (2000) 29–38.
[34] Munns R., Physiological processes limiting plant-growth in saline soils - some dogmas and hypothesis, Plant Cell Environ. 16 (1993) 15–24.
[35] Ni M., Cui D., Einstein J., Narasimhulu S., Vergara C.E., Gelvin S.B., Strength and tissue specificity of chimeric promoters derived from the octopine and mannopine synthase genes, Plant J. 7 (1995) 661–676.
[36] Parinov S., Sundaresan V., Functional genomics in *Arabidopsis:* large-scale insertional mutagenesis complements the genome sequencing project, Curr. Opin. Biotechnol. 11 (2000) 157–161.
[37] Postel S., Redesigning irrigated agriculture, in: Brown L.R., Flavin C., French H. (Eds.), State of the World 2000, W.W. Norton & Co., New York and London, 1999, pp. 39–58.
[38] Quarrie S.A., Laurie D.A., Zhu J., Lebreton C., Semikhodskii A., Steed A., Wisenboer H., Calestani C., QTL analysis to study the association between leaf size and abscisic acid accumulation in droughted rice leaves and comparison across cereals, Plant Mol. Biol. 35 (1997) 155–165.
[39] Richmond T., Somerville S., Chasing the dream: plant EST microarrays, Curr. Opin. Plant Biol. 3 (2000) 108–116.
[40] Selkov E., Galimova M., Goryanin I., Gretchkin Y., Ivanova N., Komarov Y., et al., Nucleic Acids Res. 25 (1997) 37–38.
[41] Selkov E.J., Grechkin Y., Mikhailova N., Salkov E., MPW: the Metabolic Pathways Database, Nucleic Acids Res. 26 (1998) 43–45.
[42] Shi H., Ishitani M., Kim C., Zhu J.K., The *Arabidopsis thaliana* salt tolerance gene SOS1 encodes a putative Na^+/H^+ transporter, Proc. Natl. Acad. Sci. USA 97 (2000) 6896–6901.
[43] Shinozaki K., Yamaguchi-Shinozaki K., Molecular responses to dehydration and low temperature: differ-

ences and cross-talk between two stress signaling pathways, Curr. Opin. Plant Biol. 3 (2000) 217–223.

[44] Somerville C., Somerville S., Plant functional genomics, Science 28 (1999) 380–383.

[45] The European Union Arabidopsis Genome Sequencing Consortium and The Cold Spring Harbor, Washington University in St Louis and PE Biosystems Arabidopsis Sequencing Consortium, Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*, Nature 402 (1999) 769–777.

[46] Tretheway R.N., Krotsky A.J., Willmitzer L., Metabolic profiling: a Rosetta stone for genomics?, Curr. Opin. Plant Biol. 2 (1999) 83–85.

[47] VanHal N.L., Vorst O., van Houwelingen A.M., Kok E.J., Peijnenburg A., Aharoni A., van Tunen A.J., Keijer J., The application of DNA microarrays in gene expression analysis, J. Biotechnol. 78 (2000) 271–280.

[48] Weigel D., Ahn J.H., Blázquez M.A., Borevitz J.O., Christensen S.K., Fankhauser C., Ferrándiz C., Kardailsky I., Malancharuvil E.J., Neff M.M., Nguyen J.T., Sato S., Wang Z.Y., Xia Y., Dixon R.A., Harrison M.J., Lamb C.J., Yanofsky M.F., Chory J., Activation tagging in *Arabidopsis*, Plant Physiol. 122 (2000) 1003–1013.

[49] Williams J.G.K., Construction of specific mutations in photosystem-II photosynthetic reaction center be genetic engineering methods in Synechocystis-6803, Methods Enzymol. 167 (1988) 766–778.

[50] Yale J., Bohnert H.J., Transcript expression in *Saccharomyces cerevisiae* at high salinity, J. Biol. Chem. (2001) (in press).